

Apparatus for improving disturbed speech signals

Patent Number: US5699480

Publication date: 1997-12-16

Inventor(s): MARTIN RAINER (DE)

Applicant(s): SIEMENS AG (DE)

Requested Patent: DE19524847

Application Number: US19960673151 19960701

Priority Number(s): DE19951024847 19950707

IPC Classification: G10L3/02

EC Classification: G10L21/02A1

Equivalents: JP9034496

Abstract

An apparatus for the transmission of speech signals reduces the effect of disturbances on the transmission quality. Two microphones are provided for picking up speech signals. Signal processing ensues in three partial frequency bands. The microphone signals are high-pass filtered in a lower frequency band, in the middle frequency band the signal is weighted with a scalar factor, so that this frequency band is damped during the speech pauses, and in the upper frequency band, an adaptive filter is used. At the beginning of the processing, a treble enhancement of the signals ensues, which is canceled by an inverse filter before the output of the improved signal.

Data supplied from the esp@cenet database - I2

⑯ BUNDESREPUBLIK
DEUTSCHLAND



DEUTSCHES
PATENTAMT

⑯ Patentschrift
⑯ DE 195 24 847 C 1

⑯ Int. Cl. 6:
G 10 L 7/04
G 10 L 7/02

DE 195 24 847 C 1

⑯ Aktenzeichen: 195 24 847.3-53
⑯ Anmeldetag: 7. 7. 95
⑯ Offenlegungstag: —
⑯ Veröffentlichungstag der Patenterteilung: 13. 2. 97

Innerhalb von 3 Monaten nach Veröffentlichung der Erteilung kann Einspruch erhoben werden

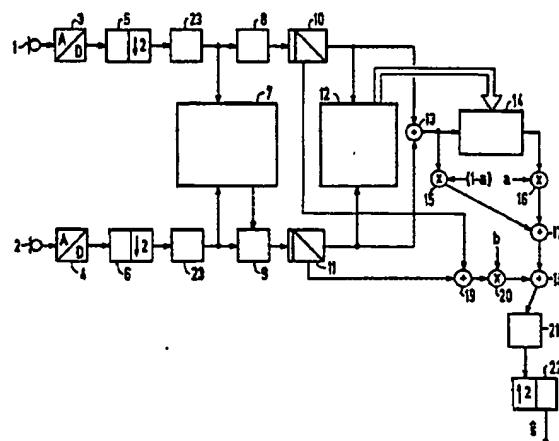
⑯ Patentinhaber:
Siemens AG, 80333 München, DE

⑯ Erfinder:
Martin, Rainer, Dipl.-Ing., 52064 Aachen, DE

⑯ Für die Beurteilung der Patentfähigkeit in Betracht gezogene Druckschriften:
DE 38 08 038 C2
DE 32 30 391 C2

⑯ Vorrichtung zur Verbesserung gestörter Sprachsignale

⑯ Es soll eine Vorrichtung zur Übertragung von Sprachsignalen geschaffen werden, bei der die Auswirkung von Störungen auf die Übertragungsqualität reduziert ist. Es sind zwei Mikrophone (1, 2) vorhanden. Die Signalverarbeitung erfolgt in drei Frequenzteilbändern. Im niedrigen Frequenzband werden die Mikrophonsignale hochpaßgefiltert, im mittleren Frequenzband wird das Signal mit einem skalaren Faktor gewichtet, so daß dieses Frequenzband während der Sprechpausen gedämpft wird. Im oberen Frequenzband wird ein adaptives Filter eingesetzt. Zu Beginn der Verarbeitung erfolgt eine Höhenanhebung der Signale, die vor der Ausgabe des verbesserten Signals durch ein inverses Filter (22) wieder rückgängig gemacht wird.



DE 195 24 847 C 1

Beschreibung

Sprachsignale können in der Medizintechnik zur Übertragung der Informationen eines Patienten benutzt werden. Insbesondere in der Computertomographie oder Magnetresonanz, bei der der Patient in einem Untersuchungsgerät liegt, kann die Kommunikation des Patienten mit dem Bedienpersonal über Mikrophone im Untersuchungsgerät erfolgen. Dabei ist es erforderlich, die Sprachsignale möglichst störungsfrei nach außen zu übertragen.

In DE 32 30 391 C2 und DE 38 08 038 C2 sind verfahren zur Signalverbesserung von gestörten Sprachsignalen beschrieben, bei denen diese Signale mit Hilfe eines einzigen Mikrofons aufgenommen werden. Nichtstationäre Störgeräusche können dabei prinzipiell nicht reduziert werden.

Der Erfindung liegt die Aufgabe zugrunde, eine Vorrichtung anzugeben, durch die nichtstationäre, gestörte Sprachsignale so weit verbessert werden, daß sich die Störungen nicht negativ auf die Informationsübertragung auswirken.

Diese Aufgabe ist erfundungsgemäß gelöst durch die Merkmale des Patentanspruches 1. Eine Weiterbildung ergibt sich aus dem Patentanspruch 2.

Die Erfindung ist nachfolgend anhand eines in der Zeichnung dargestellten Ausführungsbeispiels näher erläutert. Es zeigen:

Fig. 1 eine schematische Darstellung eines medizinischen Untersuchungsgerätes mit einer Übertragungsvorrichtung für Sprachsignale,

Fig. 2 eine Einzelheit der Übertragungsvorrichtung gemäß Fig. 1, und

Fig. 3 eine Darstellung des in Fig. 2 gezeigten Rechengliedes.

In der Fig. 1 ist ein medizinisches Gerät, z. B. ein Computertomograph, dargestellt, das ein Meßfeld aufweist, in dem ein Patient liegt. Für die Kommunikation des Patienten mit der Außenwelt sind am Gerät zwei Mikrofone 1, 2 angebracht, deren Signale über ein Filter nach außen übertragen werden.

Die Fig. 2 zeigt den Filteraufbau genauer. Den Mikrofonen 1, 2 sind A/D-Umsetzer 3, 4, Tiefpaßfilter 5, 6 zur Halbierung der Abtastrate, ein adaptiver Laufzeitausgleich 7, Übertragungsglieder 8, 9, Tiefpaß/Hochpaßfilter 10, 11 zur Frequenzbandaufteilung, ein Rechenglied 12 zur adaptiven Berechnung der Koeffizienten eines adaptiven Filters 14, das an einem Summierer 13 angeschlossen ist, Multiplizierer 15, 16, Summierglieder 17, 18, ein Summierglied 19 und ein Multiplizierer 20, ein Hochpaßfilter 21 und ein Tiefpaßfilter 22 zur Verdopplung der Abtastrate nachgeschaltet.

Der Algorithmus ist für eine Abtastrate von 8 kHz ausgelegt. Höhere Abtastraten sind bei der vorgegebenen Rechenkapazität nicht möglich und auch nicht unbedingt erforderlich, da eine Tiefpaßbegrenzung des Signals auf 3,6 kHz wegen der breitbandigen Störungen als subjektive Verbesserung des Signals empfunden wird.

Der Algorithmus umfaßt die folgenden Komponenten:

Digitaler rekursiver Tiefpaß 5 und 6. Ordnung und Abtastratenwandlung von 16 kHz auf 8 kHz. Die Abtastratenumsetzung ist erforderlich, da sich die A/D-Umsetzer 3, 4 in der vorliegenden Hardware nicht auf eine Abtastrate von 8 kHz umschalten lassen.

Automatischer Laufzeitausgleich mittels Korrelation und Maximumsuche und SNR-Detektion (SNR = Si-

gnal/Rausch-Verhältnis). Der Laufzeitausgleich der Mikrophonsignale ist auf ca. ein halbes Abtastintervall genau.

Frequenzbandaufteilung bei 800 Hz zur Verminderung tieffrequenter Geräusche. Nur das obere Frequenzband wird der adaptiven Filterung zugeführt.

Störgeräuschunterdrückung mit zwei adaptiven Filtern im Rechenglied 12, Summensignalfilter 14 und Pre-emphase-Filtern 23. Die adaptiven Filter im Rechenglied 12 werden z. B. mit dem NLMS-Algorithmus linearphasig nachgestellt. Die Koeffizientenzahl dieser Filter kann in Abhängigkeit von der Prozessorauslastung in geringen Grenzen variiert werden. Für die linearphasige Verarbeitung sind maximal 59 Koeffizienten vorgesehen. Die Koeffizienten des Summensignalfilters 14 werden spektral geglättet.

Variabile Mischung von gestörtem Eingangssignal und gefiltertem Ausgangssignal zur Verbesserung des subjektiven Eindrucks.

Dämpfung des tieffrequenten Signalzweiges mit variablem Faktor b ($0.05 \leq b \leq 0.8$).

Digitaler rekursiver Hochpaß 21 zur Unterdrückung tieffrequenter Störgeräusche. Die Grenzfrequenz liegt bei 240 Hz, die Sperrdämpfung beträgt ca. 20 dB. Die Welligkeit im Durchlaßbereich ist kleiner als 0,5 dB. Es wird vorausgesetzt, daß die analogen Hochpaßfilter der A/D-Umsetzer 3, 4 aktiv sind.

Digitaler nichtrekursiver Tiefpaß 22 der Ordnung 12–20 und Abtastratenwandlung von 8 kHz auf 16 kHz.

Für die Implementation sind folgende Punkte wesentlich:

1. Die Filterung der Mikrophonsignale mittels digitaler Hochpaßfilter ist an den Ausgang des Störunterdrückungssystems verlegt. Wegen der Bandaufteilung und der Preemphase-Filterung wird die Adaption des Störgeräuschunterdrückungsfilters nicht mehr durch tieffrequente Störanteile gestört, so daß diese Filterung auch nach der adaptiven Filterung erfolgen kann.

2. Das Signal im Tiefpaßsignalzweig wird adaptiv in Abhängigkeit des im Zuge des Laufzeitausgleichs bestimmten SNR gewichtet. Man erreicht dadurch eine zusätzliche Dämpfung des Störgeräusches in den Sprachpausen.

3. Zur weiteren Optimierung des Reststörgeräusches werden in den Sprachpausen hochfrequente Anteile durch ein Tiefpaßfilter bedämpft. Die Dämpfung wird nach den gleichen Kriterien wie die Dämpfung des tieffrequenten Signalzweiges durchgeführt.

Folgende Varianten sind möglich:

1. Variante: Das adaptive Filter 14 am Ausgang des Systems entfällt. Es werden die gefilterten Signale der adaptiven Filter im Rechenglied 12 direkt an das nachfolgende Summierglied 18 ausgegeben. Diese Variante hat den geringsten Aufwand und eine gute Sprachqualität.

2. Variante: Die im Rechenglied 12 gefilterten Signale werden zusätzlich mit dem Filter 14 gefiltert (doppelte adaptive Filterung). Diese Variante hat die höchste Störgeräuschunterdrückung aber auch die schlechteste Sprachverständlichkeit.

Die beschriebene Vorrichtung zeichnet sich insbeson-

dere dadurch aus, daß,
die Verarbeitung in drei Frequenzteilbändern vorgenommen wird im Frequenzband 0—240 Hz die Mikrophonsignale hochpaßgefiltert werden;
im Frequenzband 240—800 Hz das Signal mit einem 5
skalaren Faktor gewichtet wird, so daß dieses Frequenzband während der Sprachpausen gedämpft wird;
die skalare Gewichtung in dem Frequenzband 240—800 Hz auf der Grundlage eines geschätzten SNR 10
eingestellt wird;
im oberen Frequenzband 800 bis 3400 Hz ein adaptives Filter 14 eingesetzt wird, das durch Mittelung aus zwei linearphasig adaptierten Filtern berechnet wird, wobei zur Adaption ein entsprechender Algorithmus eingesetzt wird und die Koeffizienten spektral geglättet werden, wobei die spektrale Glättung durch die Wichtung 15
der Filterkoeffizienten des Filters 14 mit einer geeigneten Fensterfunktion erreicht wird;
zu Beginn der Verarbeitung eine Höhenanhebung der Signale erfolgt (Preemphase), die vor der Ausgabe des 20
verbesserten Signals durch ein inverses Filter wieder rückgängig gemacht wird.

Die Fig. 3 zeigt ein Ausführungsbeispiel des Rechen-gliedes 12. Die Verzögerungen T_H sind so zu wählen, daß die adaptiven Filter ein nichtkausales Wiener Filter 25
approximieren.

Patentansprüche

1. Vorrichtung zur Verbesserung gestörter Sprachsignale mit mindestens zwei Mikrofonen, wobei die von den Mikrofonen aufgenommenen Signale aus einem sprachsignal- und einem Störgeräuschanteil bestehen, dadurch gekennzeichnet, daß die Verarbeitung der Mikrophonsignale mit dem 30 Ziel einer Reduktion des Störgeräuschteils in drei Frequenzteilbändern vorgenommen wird; das niedrige Frequenzband durch ein Hochpaßfilter (21) gedämpft wird;
im mittleren Frequenzband das Signal mit einem 40
skalaren Faktor b (20) gewichtet wird, so daß dieses Frequenzband während der Sprachpausen gedämpft wird; die skalare Gewichtung (20) in dem mittleren Frequenzband auf der Grundlage eines geschätzten 45 SNR eingestellt wird; im oberen Frequenzband ein adaptives Filter (14) eingesetzt wird, das durch Mittelung aus zwei linearphasig adaptierten Filtern in einer Vorrichtung (12) berechnet wird, wobei die Koeffizienten spektral geglättet werden; zu Beginn der Verarbeitung eine Höhenanhebung 55 (23) der Signale erfolgt, die vor der Ausgabe des verbesserten Signals durch ein inverses Filter wieder rückgängig gemacht wird.
2. Vorrichtung nach Anspruch 1, bei dem das niedrige Frequenzband zwischen 0 und 240 Hz, das mittlere Frequenzband zwischen 240 und 800 Hz und das obere Frequenzband zwischen 800 und 3400 Hz liegt.

60

Hierzu 2 Seite(n) Zeichnungen

- Leerseite -

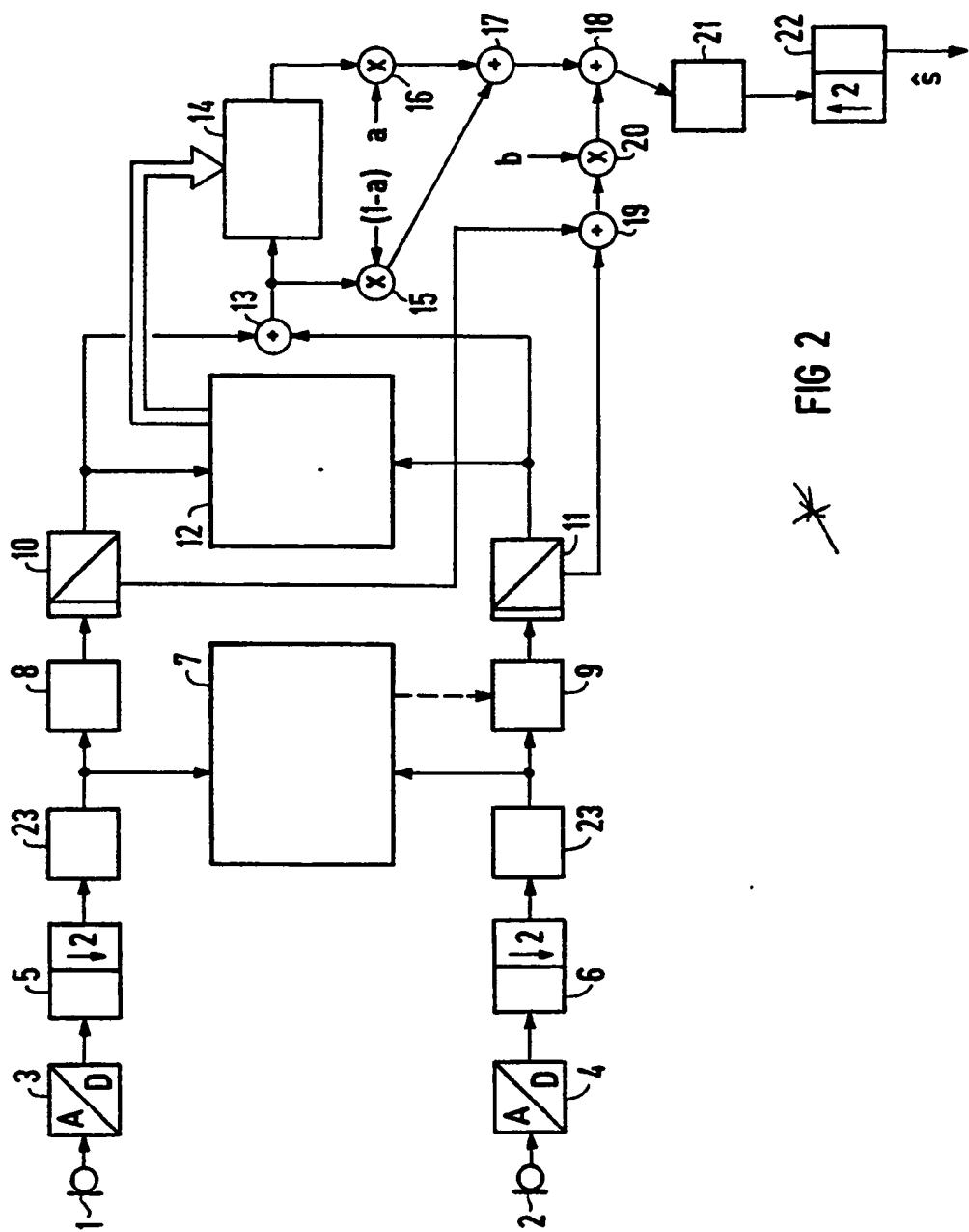


FIG 2



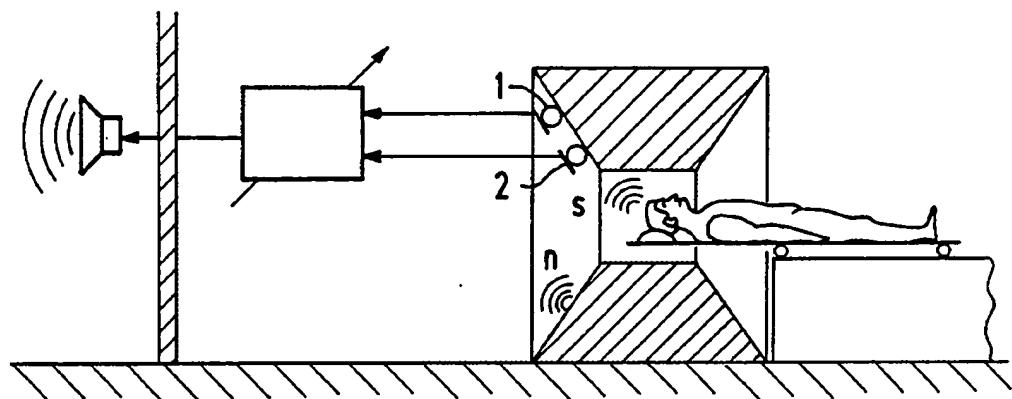


FIG 1

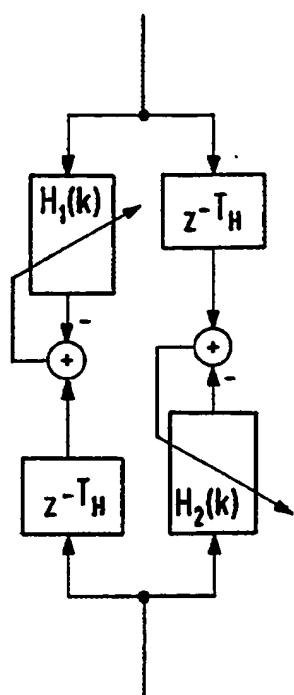


FIG 3